DOCUMENT RESUME

ED 202 881 TM 810 303

AUTHOR	Cummings, Oliver W.
TITLE	A Student-Centered Interpretation Technique Using Individual Performance on Skills Objectives of the Iowa Tests of Basic Skills.
SPONS AGENCY	National Inst. of Education (ED), Washington, D.C.
PUB DATE	Apr 81
GRANT	NIE-G-80-0084
NOTE	20p.; Paper presented at the Annual Meeting of the American Educational Research Association (65th, Los Angeles, CA, April 13-17, 1981).
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Attitude Measures; *Cluster Analysis; Grade 4; Grade 5; Grade 6; Intermediate Grades; *Program Effectiveness; *Student Attitudes; *Teacher Attitudes; *Test Interpretation
IDENTIFIERS	*Iowa Tests of Basic Skills

ABSTRACT

Research on an interpretation technique (IT) for using the Pupil Item Response Record (PIRR) of the Iowa Tests of Basic Skills (ITBS) to actively involve students in their test interpretation was presented. The major concern of the study was to determine whether the IT had any impact on attitudes toward or knowledge about the ITBS. Students in grades 4-6 used their own PIRR and Skill Summary Sheet (SSS), a listing of the major skills tested, to do their summarizations. A substantially different form of instrument was administered to 52 teachers in those grades. Preliminary results indicated: (1) a significant difference in knowledge found between students who had and those who had not been through the interpretation session; (2) a small, persistent decline, on average, in the students' attitudes of "caring about the test" over the three grades; (3) students' attitudes, at all grade levels, that most of the items covered material they had studied; (4) an attitude by students who participated in the IT that they had done better on the ITBS than students who did not participate; and (5) no immediate impact on either teacher attitudes or on teacher knowledge about the ITBS. (RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED202881

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

A STUDENT-CENTERED INTERPRETATION TECHNIQUE USING INDIVIDUAL PERFORMANCE ON SKILLS OBJECTIVES OF THE IOWA TESTS OF BASIC SKILLS

A paper presented at the Annual Meeting of the
American Educational Research Association
Los Angeles, California, April, 1981

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
O. W. Cummings
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Oliver W. Cummings
Grant Wood Area Education Agency
Cedar Rapids, Iowa

This research is supported, in part, by the National Institute of Education under Grant Number NIE-G-80-0084. Any opinions, findings, and conclusions or recommendations expressed are those of the author and do not necessarily reflect the views of the Institute or the Department of Education.

TM 810 303

A STUDENT-CENTERED INTERPRETATION TECHNIQUE USING INDIVIDUAL PERFORMANCE ON SKILLS OBJECTIVES OF THE IOWA TESTS OF BASIC SKILLS

A paper presented at the annual meeting of the
American Educational Research Association
Los Angeles, California, April, 1981

Over the past several years, publishers of the most widely used standardized achievement tests have developed scoring services which report raw scores or the percent of items correct on item groupings representing the subskills measured on the tests.[1] These reports, variously known as Cluster Analysis, Individual Test Record, Pupil Item Response Record, etc., are claimed to be useful in determining specific student strengths and weaknesses (Finch, et al., 1976, p. 42; Prescott, et al., 1978, p. 33) and for selective followup (Hieronymus, et al., 1979, p. 31). The development and subsequent acceptance of score reports of this type are based in a body of literature which calls for specific interpretation of item responses in working with and planning for students (Bradley, 1978; Ebel, 1972; Rudman, 1977). The encouragement of such interpretations can be generally summarized by Ebel's (1972) statement:

> "Any achievement tests can provide 'diagnostic' information of value to the individual pupil if (s)/he is told which items (s)/he missed. Then if (s)/he chooses to do so, (s)/he can, with the teacher's help, correct the mistakes or misconceptions that led him(her) astray. Highly specific 'diagnosis' and 'remediation' of this sort can be effective and ought to be encouraged." (p. 478)

---

[1] The eight most widely used programs, Basic Skills Assessment, California Achievement Test, Comprehensive Tests of Basic Skills, Educational Development Series, Iowa Tests of Basic Skills, Metropolitan Achievement Test, SRA Achievement Series, and Stanford Achievement Test, were each found to provide such reports.

This paper presents research on a technique for using the Pupil Item Response Record of the Iowa Tests of Basic Skills to actively involve students in their test interpretation. The process, developed as a method for teachers to use in communicating with students about their performance on the test, was studied during the 1980-81 school year to determine both student and teacher reaction to the interpretation sessions. In addition, the impact of such an interpretation technique upon the attitudes toward tests and knowledge about the test for both students and teachers was studied.

The Interpretation Technique

Interpretations were conducted in classroom sized groups and required approximately 40 to 50 minutes per group. The interpretation process involved several steps leading to each student summarizing his/her own performance on each of approximately 60 skills identified on the Iowa Tests of Basic Skills. The students used their own scoring service report form (Pupil Item Response Record) and a Skill Summary Sheet, constructed for use in this project, to do the summarization.

The Pupil Item Response Record provides information on each student's answer for each question on the test. In addition to identifying information, scale scores, and other information, the student's response to each item (correct, incorrect, or omit) and the skill measured by each item are provided on the report.

The Skill Summary Sheet is a listing of the major skills tested. For each major skill tested, three broad categories of performance have been defined: 1) Possible problem area, 2) More information needed; and 3) Satisfactory progress likely. The three categories have been defined on the sheet according to raw score performance, adjusted for differences in mean difficulty of the set of items used to test each skill.

The interpretation of results involves a relatively simple, six-step process.

Step 1: Students are introduced to the test interpretation process. They are given a general explanation of the meaning of the tests and how the results can be used to improve instruction. A further explanation that each of the tests (i.e., Mathematics Concepts, Language Usage, etc.) consists of questions that measure several specific skills is included.

Step 2: Students are given their own Pupil Item Response Record and a copy of the appropriate Skill Summary Sheet.

Step 3: Students are taught how to read and understand the Pupil Item Response Record, and what the parts of the form mean.

Step 4: Students are taught how to record their own performance for each major skill area on the Skill Summary Sheet. They are instructed to count the number of pluses (correct answers) for each skill and to circle the corresponding number on the Skill Summary Sheet.

Step 5: Students complete the Skill Summary Sheets, as the teacher or counselor circulates around the group providing assistance as needed.

Step 6: Students are given instruction on the meaning of the Skill Summary Sheet. They are encouraged to discuss their results or questions with their counselor or teacher to determine the significance of any particular high or low performance. General comments about the limitations of the tests and how to keep the results in perspective are also made.

For a more thorough discussion of the interpretation technique and an example, see Cummings' (1981) article, Student-Centered Test Interpretation: An Active Technique.

PRELIMINARY RESEARCH RESULTS

Participants

During the fall and winter of the 1980-81 school year data were collected on student and teacher attitudes toward testing, and on their knowledge about the Iowa Tests of Basic Skills. These data were collected for 1,116 fourth-, fifth-, and sixth-grade students, and for 52 teachers in those grades. The sample included 52 classroom groups from six small city and rural school districts in eastern Iowa. The districts ranged in size from 106 to 3,316 students.

The Study Design

Since the major concern of the study was to determine whether the interpretation technique had any immediate impact on attitude or knowledge, a conservative design was used. Schools were solicited for participation and were provided interpretation services in return for a commitment of approximately 20 additional minutes for each class group to spend in responding to the questionnaire forms used in the study. Classrooms were randomly assigned to either experimental or control groups. Both students and their teachers were assigned on the same basis to be either control or experimental subjects.

Control groups responded to the attitude and knowledge questions immediately prior to the interpretation session, and experimental groups responded to the questions immediately following the interpretation session. Thus, in each case the participants in the study were in schools that were at least willing to engage in a substantial effort toward test use and interpretation to students. It was believed that this design would more accurately assess the impact of the interpretation session than one which could have included schools which were not giving the tests or in which the general attitude toward the use of the test results was not conducive

to major efforts at interpretation. This design, then, was one in which pre-existing attitudes toward the test and knowledge about the Iowa Tests of Basic Skills, in particular, were likely to be relatively high, and in which observed differences between control and experimental groups were likely to be directly related to the interpretation process.

In addition to the knowledge and attitude questions, both students and teachers were asked to complete a short evaluation form following the interpretation session. The preliminary results of these attitude, knowledge, and evaluation assessments are presented below.

Knowledge Instruments and Results for Students and Teachers

The knowledge instrument used for students in the study was a 14-item true/false test, covering knowledge about the uses, purposes, and structure of the Iowa Tests of Basic Skills. While the analysis of the test characteristics has not been completed for this preliminary report, certain data are available. The range of p-values for the test is from approximately .36 to approximately .93, with a mean p-value across the total group of 1,116 students of .688. Assuming a mean point biserial correlation of a single item with total test score of approximately .40, a reasonably conservative assumption for a test of limited scope (Thorndike, 1971, p. 66), the estimated reliability of the test would be about .60.

The impact of the interpretation sessions on the knowledge of students about the Iowa Tests of Basic Skills was tested using a t-test of differences in mean scores on the knowledge instrument between the experimental ($n = 457$, $\bar{x} = 9.75$) and control ($n = 659$, $\bar{x} = 9.38$) groups. A t-value of -2.84 ($p = .0046$) was obtained, with the higher mean score in favor of the experimental group on the knowledge instrument.

For the teacher groups, a substantially different form of the knowledge instrument was administered. No item or test statistics on this

instrument were available for this preliminary report. However, the results of a t-test of differences in mean performance between teachers in the control ($n = 27$, $\bar{x} = 9.52$) and experimental ($n = 29$, $\bar{x} = 9.38$) groups yielded no significant difference between the two groups ($t = .26$, $p = .7964$).

Attitude Items and Results for Students

Analysis of variance procedures were carried out for seven attitude/opinion questions related specifically to the Iowa Tests of Basic Skills for students in the experimental and control groups. The questions, designated as A1 through A7 here, were as follows:

A1: How well do you think you did on the Iowa Tests of Basic Skills this year?

A. Quite high
B. Above average
C. Average
D. Below average
E. Quite low

A2: In general, how do you feel about the Iowa Tests of Basic Skills?

A. I really like them
B. I like them
C. I don't care one way or the other
D. I hate them
E. I really hate them

A3: How hard do you think tests like the Iowa Tests of Basic Skills are?

A. Very hard
B. Hard
C. Medium
D. Easy
E. Very easy

A4: How nervous do you feel before you take a test like the Iowa Tests of Basic Skills?

A. Extremely nervous
B. Very nervous
C. Nervous
D. Just a little nervous
E. Not at all nervous

A5: How many questions on the Iowa Tests of Basic Skills cover things you have studied in school?

A. All of them
B. Most of them
C. Some of them
D. Only a few of them
E. None of them

A6: How much do you think you know about the tests on the Iowa Tests of Basic Skills?

A. A lot
B. Quite a bit
C. A little
D. Not much at all
E. Nothing

A7: How useful are the Iowa Tests of Basic Skills results to you?

A. Extremely useful
B. Very useful
C. Useful
D. Not useful
E. Not at all useful

The results of the attitude questions are reported here by grade and group. The number of students for each cell and for the totals are presented in Table 1, and remained reasonably constant over the seven questions.

Table 1. Number of Students by Grade and Group.

	Grade 4	Grade 5	Grade 6	Total
Experimental	99	223	134	456
Control	111	291	252	654
Total	210	514	386	

The mean scores for the attitude questions are presented in Table 2. These means are provided by grade, group, and for totals. Response position A on an item was scored 5 points, and response position E was scored 1 point.

Table 2. Means by Grade, Group, and Total.

QUESTION	GROUP	GRADE 4	GRADE 5	GRADE 6	TOTAL
A1	Experimental	3.50	3.37	3.31	3.38
	Control	3.25	3.26	3.29	3.27
	Total	3.37	3.31	3.30	
A2	Experimental	2.89	3.21	2.89	3.04
	Control	3.49	3.09	2.93	3.20
	Total	3.20	3.14	2.92	
A3	Experimental	3.50	3.17	3.35	3.30
	Control	3.26	3.26	3.26	3.26
	Total	3.37	3.22	3.29	
A4	Experimental	3.01	2.68	2.80	2.79
	Control	2.86	2.74	2.73	2.76
	Total	2.93	2.72	2.75	
A5	Experimental	3.52	3.68	2.71	3.65
	Control	3.70	3.70	3.67	3.69
	Total	3.62	3.69	3.69	
A6	Experimental	3.58	3.50	3.59	3.54
	Control	3.56	3.56	3.52	3.54
	Total	3.57	3.53	3.54	
A7	Experimental	3.58	3.63	3.68	3.64
	Control	3.68	3.70	3.51	3.62
	Total	3.64	3.67	3.57	

A test of significance yielded a group main effect difference for question A1 ($F_{1, 1104} = 4.84$, $p = .03$). The experimental group indicated that they thought they had done better on the Iowa Tests of Basic Skills than the control group. There were no significant grade differences nor grade x group interactions.

For question A2, which asked about how well students like the Iowa Tests of Basic Skills, no significant difference was found between the experimental and control groups. However, there was a significant main effect for grade ($F_{2, 1101} = 7.36$, $p = .001$) and for group x grade interaction ($F_{2, 1101} = 8.63$, $p = .0002$).

When asked how hard they thought the tests were, question A3, no significant difference was found between experimental and control groups. There were, as with question A2, significant main effects for grade ($F_{2, 1103} = 3.12$, $p = .04$) and for group x grade interaction ($F_{2, 1103} = 3.94$, $p = .02$).

For the remaining four attitude questions no significant differences were found.

Attitude Items and Results for Teachers

Analysis using t-test procedures were carried out for the following eight attitude/opinion questions asked of teachers:

A. How relevant are the results of the ITBS to your work with students?

1. Not at all relevant
2. Not very relevant
3. Somewhat relevant
4. Very relevant
5. Extremely relevant

B. How useful are the results of the ITBS in identifying strong or weak points in the curriculum?

1. Not at all useful
2. Minimally useful
3. Useful to some extent
4. Useful to a great extent
5. Useful to a very great extent

C. How useful are the results of the ITBS in discussing future instructional plans with individual students?

1. Not at all useful
2. Minimally useful
3. Useful to some extent
4. Useful to a great extent
5. Useful to a very great extent

D. How closely do the skills tested on the ITBS match the skills in the curriculum you actually teach?

1. Very high match
2. High match
3. Medium match
4. Low match
5. Very low match

E. To what extent do you think the results of the ITBS can be used for improving students' understanding of their specific strengths and weaknesses?

1. Not at all
2. To a minimal extent
3. To some extent
4. To a great extent
5. To a very great extent

F. How useful are the results of the ITBS in helping parents better understand the strengths and limitations of their child?

1. Not at all useful
2. Minimally useful
3. Useful to some extent
4. Useful to a great extent
5. Useful to a very great extent

G. How well informed do you consider yourself to be about the ITBS?

1. Not informed
2. Minimally informed
3. Informed
4. Well informed
5. Extremely well informed

H. How would you rank the overall quality of the ITBS as compared to other standardized tests of its type?

1. One of the best
2. Above average
3. About the same as others
4. Below average
5. One of the worst

Table 3 presents means and t-tests for the teacher attitude/opinion questions. For purposes of the analysis, questions B, C, E, and F were combined to create a scale for "usefulness." Each item was also handled independently. No significant differences were found on any of the attitude/opinion questions for experimental versus control teachers.

Table 3. Means and t-tests for Teacher Attitude/Opinion.

	EXPERIMENTAL			CONTROL			
Variable	n	Mean	SD	n	Mean	SD	t	p>|t|
Usefulness	29	16.17	3.48	27	16.78	2.06	.79	.43
A	29	3.10	.94	27	3.33	.62	1.09	.28
B	29	3.45	.69	27	3.56	.58	.63	.53
C	29	2.97	.98	27	3.04	.59	.33	.74
D	29	2.79	.90	25	2.72	.61	-.34	.73
E	29	3.55	.91	27	3.52	.70	-.15	.88
F	29	3.10	.86	27	3.33	.68	1.10	.27
G	28	3.43	.79	27	3.15	.72	-1.38	.18
H	28	1.82	.77	27	2.00	.68	.91	.37

Evaluations of the Interpretation Sessions

In each of the 52 separate interpretation sessions that was conducted for this study, students and teachers were asked to respond to a short set of questions to evaluate the session in which they were involved. The results of these evaluations are presented descriptively below.

Students were asked to respond to three questions:

E1: Do you think that knowing your stong and weak areas will help you learn better?

A. Yes
B. No
C. Not sure

E2: Do you think the skill session was confusing?

A. Yes
B. No
C. Not sure

E3: Do you think the skill session was interesting?

A. Yes
B. No
C. Not sure

Table 4 presents a summary of responses by frequency and percent for the three evaluation questions asked of students.

Table 4. Summary of Responses for Student Evaluation Questions.

	Yes		No		Not Sure	
Question	f	%	f	%	f	%
E1	811	77.1	69	6.6	170	16.2
E2	215	20.4	577	54.8	261	24.8
E3	567	54.3	299	28.6	179	17.1

Teachers were asked to respond to five evaluation questions:

A. How difficult do you think the interpretation was for your students?

1. Too difficult
2. About right
3. Too easy

B. How would you rate student interest in the interpretation session?

1. Very interested
2. Somewhat interested
3. Neutral
4. Somewhat bored
5. Very bored
. Don't know

C. Do you think the interpretation session will positively affect the students' test taking attitudes?

1. Yes
2. No
3. Not sure

D. Do you think that the interpretation session and follow-up on it will result in improved teaching/learning?

1. Yes
2. No
3. Not sure

E. Do you think that this type of interpretation session is worth continuing next year?

1. Yes
2. No
3. Not sure

In response to question A, 51 teachers (100 percent of those responding to the question) indicated that the difficulty level of the session for their students was about right. The response to a judgment about the interest level of the students was slightly more variable. Thirty-five teachers (67 percent) said their students were "very interested," 15 teachers (29 percent) said their students were "somewhat interested," and two teachers (4 percent) said their students were "neutral," in terms of interest. No teachers said their students were either somewhat or very bored with the session.

Questions C and D represented an attempt to find out what teachers perceived to be the potential impact of the sessions on two important goals of the sessions. For question C, 29 teachers (56 percent) believed the

interpretation session would positively affect students' test taking attitudes. The remaining 23 teachers were either not sure (21 teachers--40 percent) or felt the session would not positively affect student attitudes (two teachers--4 percent).

For question D, 34 teachers (65 percent) believed the session and follow-up on it would result in better teaching or learning. Seventeen teachers (33 percent) were not sure, and one teacher (2 percent) said it would not result in better teaching or learning.

No teachers indicated that the interpretation sessions should be discontinued for next year. Forty-nine teachers (94 percent) said the sessions are worth continuing next year, and three teachers (6 percent) said they were not sure whether the sessions should be continued.

DISCUSSION OF PRELIMINARY RESULTS

Results for Students

One finding in this study which holds the potential for significant impact on students' feelings and attitudes about the Iowa Tests of Basic Skills is the significant difference in knowledge found between students who had and those who had not been through the interpretation session. Cormany's (1974) study of attitudes toward standardized testing concluded that persons who felt they were well informed about the subject had more positive attitudes about it. If the increase in knowledge about the test, generated through the interpretation session, leads to feelings of being well informed (or better informed) about the test, then general attitudes toward the test may be improved over the long run.

The emphasis of this discussion of attitude change resulting from greater knowledge, however, must be on the long term potential effects, since no group differences were found for the short term effects of the

interpretation sessions; the groups did not differ on opinion questions related to how students feel about the test, how much they think they know about it, or how nervous they feel before entering the testing situation.

In general, student's feelings about the test (A2) centered around the "I don't care one way or the other" option. However, there appears, on average, to be a small, persistent decline in this student attitude over the three grades. Fourth and fifth graders tended to score to the positive side of this neutral position, and sixth graders to the negative side. This shift in attitude across grades may account for the significant interaction and grade effects found in the results.

While there were no differences between the groups in their opinions of the difficulty level of the test, the judgments of the students corresponds well to the actual difficulty of the test. The students judged the test to be slightly to the hard side of medium in difficulty level. This evaluation of the test is consistent with the mean item difficulties reported for the Iowa Tests of Basic Skills, which are generally more difficult than teacher made tests.

Students in both groups, at all grade levels were somewhat nervous when faced with taking the test, but felt that some to most of the items covered things they had studied. Further, they felt that the results of the test were useful to very useful and that they knew more than a little about the test itself.

The one opinion item which appeared to be directly affected by the interpretation process concerned how well the students thought they had done on the Iowa Tests of Basic Skills. Those students who participated in the interpretation process felt they had done better on the test than those who did not participate (see frequency distributions in Figure 1).

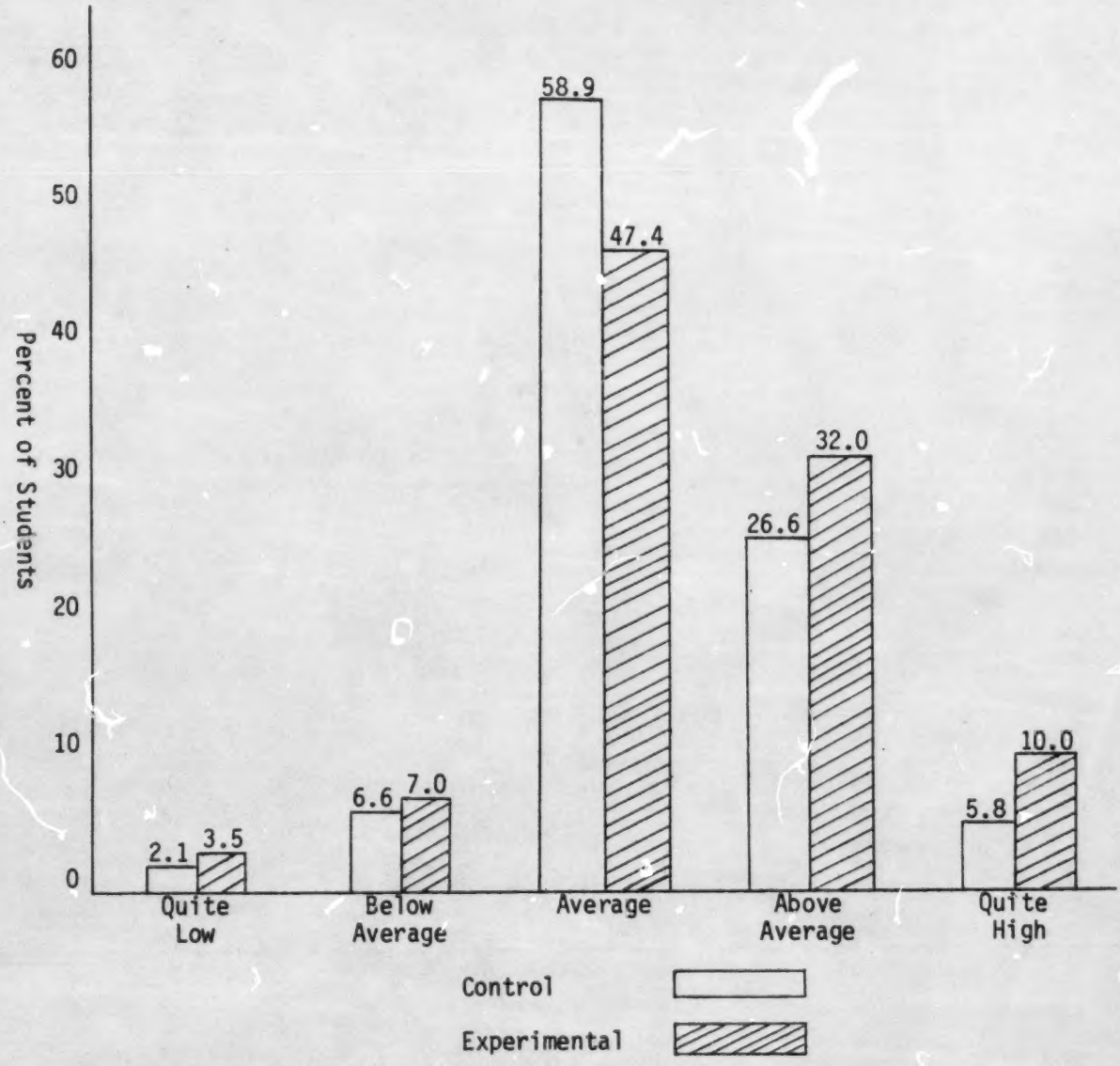

Figure 1. Frequency distributions of the experimental and control group for A2 (How well do you think you did on the Iowa Tests of Basic Skills this year?).

The results are particularly relevant in view of the frequent criticism that standardized tests may damage students' self-image. Figure 1 shows that student self-ratings, in general, clustered around the average rating with some skewness on the above average side. If the criticism were valid, the distributions would be skewed in the opposite direction. The findings further suggest that if test results are not interpreted with active

student participation, students tend to rate themselves lower on the test and have a lower self-image of their abilities to achieve in school.

In summary, the interpretation process had an immediate impact on student knowledge about the Iowa Tests of Basic Skills and on the students' views about how well they had performed on the test. Other attitudes and opinions about the test were not immediately impacted by the interpretation process. Further research on the long term effects on attitudes is needed.

Results for Teachers

The interpretation process had no immediate impact on either teacher attitudes/opinions or on teacher knowledge about the test, as assessed for the study. However, the evaluations of the interpretation sessions by teachers indicate that the sessions were positively received, were thought by most teachers to have the potential for positive effects in both future testing and teaching situations, and were considered by almost all teachers (94 percent) to be worth continuing next year.

REFERENCES

Bradley, R. W. Person-referenced test interpretation: a learning process. Measurement and Evaluation in Guidance, 1978, 10, 201-210.

Cormany, R. B. Faculty attitudes toward standardized testing. Measurement and Evaluation in Guidance, 1974, 7, 188-194.

Cummings, O. W. Student-centered test interpretation: an active technique. The School Counselor, In press.

Ebel, R. L. Essentials of Educational Measurement. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1972.

Finch, F. L. and others. Comprehensive Tests of Basic Skills: Test Coordinator's Handbook. Monterey, California: CTB/McGraw-Hill, 1976.

Hieronymus, A. N.; Lindquist, E. F.; Hoover, H. D.; and others. Iowa Tests of Basic Skills: Teachers Guide for Administration, Interpretation, and Use (Levels: 9-14). Boston: Houghton-Mifflin Company, 1979.

Prescott, G. A.; Balow, I. H.; Hogan, T. P.; and Farr, R. C. Metropolitan Achievement Tests: Teacher's Manual for Administering and Interpreting. New York: The Psychological Corporation, 1978.

Rudman, H. C. The standardized test flap. Phi Delta Kappan, 1977, November, 179-185.

Thorndike, R. L. Educational Measurement (2nd Ed.). Washington, D. C.: American Council on Education, 1971.